

## **VIDEO SEGMENTATION FOR MOVING OBJECT**

### **DETECTION & TRACKING- A REVIEW**

**PRIYA GUPTA, NIVEDITA KUMARI & SHIKHA GUPTA**

Department of Computer Science and Engineering, Government Engineering College,  
Ajmer, Rajasthan, India

#### **ABSTRACT**

The objective of this paper is to review the contemporary video segmentation & tracking methods, organize them into different categories, and identify new leanings. Object detection and tracking, in general, is a stimulating problem. The detection of moving object is important in many tasks, such as video surveillance & the moving object tracking. The design of a system of video surveillance is directed on the automatic identification of the events of interest, particularly on tracking & classification of moving objects. Complications in tracking objects can arise due to unexpected object motion, the changing appearance patterns of object & the scene, non-rigid object edifices, object-to-object and to scene obstructions, and camera motion. Tracking is typically performed in context of the higher-level applications that need the location &/or shape of object in every frame. Classically, the assumptions are made to restrain the tracking problem in context of particular application. In this examination, we classify the tracking approaches on the basis of object & the motion representations used; afford thorough descriptions of representative methods in each class. Furthermore, we deliberate the important concerns connected to the racking containing the use of suitable image features, the selection of motion representations, and detection of objects.

**KEYWORDS:** Object Detection, The Video Segmentation, Video Surveillance, Object Tracking Survey, Review

#### **1. INTRODUCTION**

Object tracking is an important task within the field of computer vision. The proliferation of high-powered computers, the obtain ability of the high quality & low-cost video cameras, & increasing essential for the automated video analysis has produced the great deal of interest in the object tracking algorithms. There are mainly 3 key stages in video analysis: the detection of interesting moving objects, the tracking of such objects from frame to the frame, & analysis of the object tracks to identify their behavior. Consequently, the use of the object tracking is relevant in tasks of:

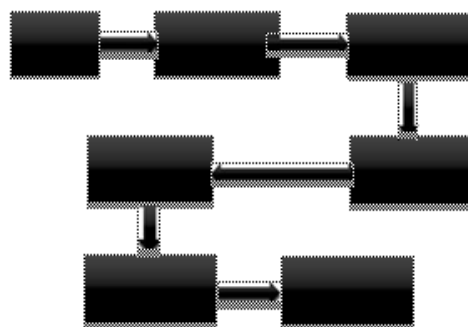
- Motion-based recognition, that is, human identification based on gait, automatic object detection, etc.;
- Automated surveillance, namely, monitoring the scene to identify the suspicious activities suspect events;
- Video indexing, that is, the automatic explanation & recovery of videos in the multimedia databases;
- Human-computer interaction, that is, the signal recognition, eye gaze tracking for the data input to computers, etc.;
- Traffic monitoring, that is, real-time gathering of traffic statistics to direct traffic flow.
- Vehicle navigation, which is, the video-based path planning& difficulty avoidance abilities.

In its modest form, tracking can be well-defined as a problem of guessing the trajectory of an object in image plane as it changes around a scene. In other words, a follower assigns reliable labels to tracked objects in dissimilar frames of video. Furthermore, depending on tracking domain, the tracker can also offer the object-centric info, like the orientation, area, /shape of an object. The tracking objects can be multiple due to:

- Loss of info produced by the projection of 3D world on a 2D image,
- Noise in images,
- Complex object motion,
- Non-rigid or articulated nature of objects,
- Partial and full object occlusions,
- Complex object shapes,
- Scene illumination changes, and
- Real-time processing necessities.

One can make simpler tracking by magnificent constraints on motion &/or entrance of the objects. For e.g., nearly all tracking algorithms accept that the object motion is flat with no sudden changes. One can additional constrain the object motion to be of endless velocity endless acceleration based on the priori info. The prior knowledge about the no. & size of objects, / the object entrance & shape, can also be used to make simpler the problem.

The huge no. of tracking approaches have been proposed which effort to answer these queries for the variation of scenarios. The goal of survey is to group tracking approaches into the wide groups & deliver the complete explanations of the representative approaches in each category. We desire to give the readers, who need a tracker for a definite application, the capability to select the most appropriate tracking algorithm for their specific requirements. Furthermore, we aim to recognize novel trends & ideas in tracking community & a hope to offer insight for the improvement of novel tracking approaches, the basic framework of the moving object detection for the video surveillance is revealed in figure below.



**Figure 1: Framework for Basic Video Object Detection System**

There has been a growing research interest in video image segmentation over the past decade and towards this end, a wide variety of methodologies have been developed [53]-[56].

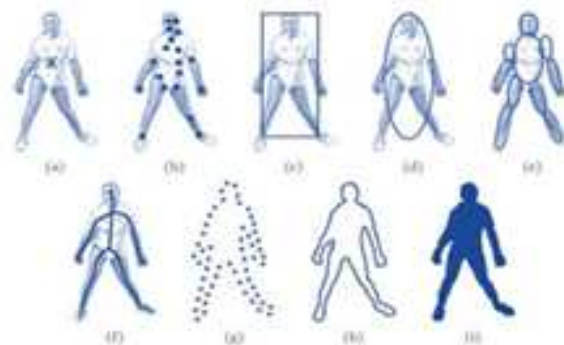
The segmentation of video methodologies have broadly used the stochastic image models, mainly the Markov Random Field (MRF) model, as a model for the video sequences [57]-[59]. The MRF model has shown to be an actual stochastic model for the segmentation of image [60]-[62] because of its quality to the model context dependent things like the image pixels & associated features. In the Video segmentation, also the spatial modeling & constraints, the temporal constraints are too added to develop the spatio-temporal schemes of image segmentation.

An algorithm of adaptive clustering has been described [57] where the temporal constraints & the temporal local density have been accepted for the smooth transition of the segmentation from frame to frame. The segmentation of spatio-temporal has too been functional to the image categorizations [63] with the different techniques of filtering. The extraction of moving object & the tracking of same has been attained in the spatio-temporal framework [64] with the Genetic algorithm helping as the optimization tool for the image segmentation.

Newly, the MRF model has been used to the model spatial individuals in each frame [64] & the Distributed Genetic algorithm (DGA) has been used to acquire the segmentation. Improved version of the DGA has been proposed [58] to acquire the segmentation of video sequences in the spatio-temporal framework. Moreover, the video segmentation & the foreground subtraction has been attained using spatio-temporal notion [65]- [66] where the spatial model is Gibbs Markov Random Field & the sequential changes are modelled by the combination of Gaussian distributions. Currently, the algorithm of automatic segmentation of the foreground objects in video sequence segmentation has been proposed [67].

### Issues in the Building an Object Tracker

In the tracking scenario, an object can be well-defined as anything that is of attention for the further analysis. For example, fish inside an aquarium, v boats on the sea, planes in the air, vehicles on a road, bubbles in the water, / people walking on a road are the set of objects that may be significant to track in a detailed domain. Objects can be signified by their shapes & appearances.

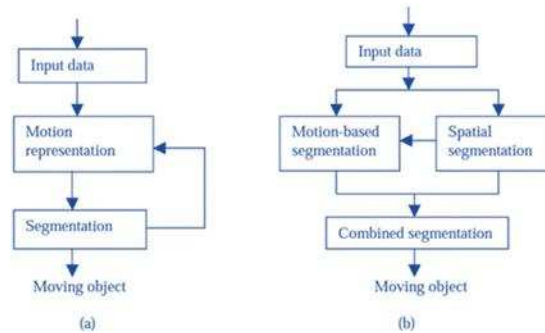


**Figure 2: An Object Representation. (A) The Centroid, (B) Multiple Points, (C) A Rectangular Patch (D) Elliptical Patch, (E) The Part-Based Multiple Patches, (F) An Object Skeleton, (G) Entire Object Contour, (H) The Control Points on an Object Contour, (I) An Object Silhouette**

Choosing the right topographies plays a serious role in the tracking. In general, the most wanted property of the visual feature is its individuality so that an object can be effortlessly distinguished in feature space. The selection of feature is carefully related to object representation. For e.g., the color is used as a feature for the histogram-based entrance

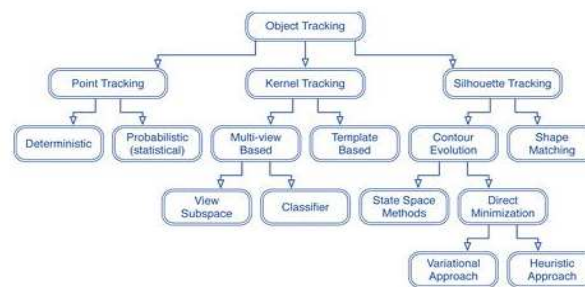
representations, while for the contour-based representation, object edges are typically used as a features. In overall, many algorithms of tracking use a mixture of these features.

The algorithms of motion-based segmentation can be categorized whichever based on their motion representations based on their clustering standards. Meanwhile the motion representation plays such a vital role in the motion segmentation that motion segmentation methods usually focus on design of motion approximation algorithms. Consequently, the motion segmentation are the best identified & well-known by motion representation it adopts. Inside each subgroup recognized by its motion representation, the approaches are illustrious by their clustering criteria.



**Figure 3: (a) Basic Motion-Based Segmentation & (b) Simplified Spatio-Temporal Segmentation**

Every tracking technique needs an object detection mechanism whichever in every frame / when the object 1st appears in video. A common method for the object detection is to use info in the single frame. Though, some object detection approaches make use of temporal info computed from an arrangement of the frames to decrease the no. of false detections. This temporal info is typically in a form of the frame differencing, which highlights altering regions in the consecutive frames. Given object regions in the image, it is then the task of tracker to complete the object correspondence from 1 frame to next to produce the tracks.



**Figure 4: Classification of Tracking Methods for Moving Object Detection in Videos**


The main attributes of such tracking algorithms can be summarized as follows.

- **Feature-Based Dense-Based:** In the feature-based approaches, the objects are characterized by the limited no. of points such as the corners the salient points, however dense approaches compute the pixel-wise motion.
- **Occlusions:** It is Capability to Deal with the Occlusions.
- **Multiple Objects:** It Is Capability to Compact with More than 1 Object in Scene.

- **Spatial Continuity:** It is Capacity to Adventure the Spatial Continuity.
- **Temporary Stopping:** It is ability to contract with the temporary stop of objects.
- **Robustness:** It is ability to contract with the noisy images (in the case of feature based approaches it is the location of point to be pretentious by noise but not data association).
- **Sequentially:** It is ability to work increasingly, this means for e.g. that the algorithm is capable to achievement info that was not current at the beginning of sequence.
- **Missing Data:** It is capacity to contract with the missing data.
- **Non-Rigid Object:** It is capacity to deal with the non-rigid objects.
- **Camera Model:** If it is essential, which camera model is used (the orthographic, para-perspective perspective).

## Structure of Assessment

The edifice step of this paper is as follows. The Introductory

Section ends with a brief introduction of Video Segmentation, Object Detection and tracking and its necessity in the field of video surveillance. The part  in introduction shows a brief explanation about principle and issues in developing an object tracker.

In Section II, we explain a General review of Mean-shift and Optical Flow based Approaches in video segmentation.

In section III we address the review on recent researches in traditional Active Contour Model, Layer and Region based Approaches, in which we have taken a review on video segmentation and object tracking techniques based on these approaches.

Section IV addresses the Stochastic and Statical Approaches, including Markov random field, MAP, PF and Expectation maximization algorithm.

Section V explains Deterministic and feature based Approaches including edge & patch based features, skeleton based approaches.

Section VI elucidates Image Difference and Subtraction based Approaches including the temporal difference & image subtraction methods.

Section VII shows the Supervised and un-supervised Learning Based Approaches. And a general conclusion of the paper, discussion regarding review is presented in Section VIII and after that a tabular comparison of different researches reviewed in previous sections are enlisted.

## 2. MEAN-SHIFT AND OPTICAL FLOW BASED APPROACHES

Mean-Shift (MS) is a clustering technique in the image segmentation for a joint spatial-color space. Mean-shift segmentation technique is used to examine the complex multi- modal feature space & the identification of the feature clusters. It is the non-parametric method. It's the Region of Interest's (ROI) size & the shape parameters are only allowed parameters on the mean-shift process, such as, multivariate density kernel estimator. A 2-step arrangement of the

discontinuity preserving filtering & mean-shift clustering is used for the mean-shift image [68, 69]. The MS algorithm is modified with the huge no. of hypothesized cluster centres casually selected from the data of given image. The MS algorithm goals at discovery of nearest stationary point of original density function of the data &, thus, its usefulness in noticing the modes of density. For this determination, each cluster center is stimulated to mean of data dishonest inside the multidimensional ellipsoid focused on the cluster center [70]. For the meantime, algorithm builds up the vector, which is definite by old & the novel cluster centres. This vector is named as the MS vector & computed iteratively until cluster centres do not alteration their positions. Some clusters may get combined during the MS iterations [71, 70, & 68]. The MS-based image segmentation (within the sense of video frame's spatial segmentation) is a direct extension of incoherence preservative smoothing algorithm. After the close modes are trimmed as in generic feature space analysis technique [68], each pixel is connected with an important mode of joint domain density situated in its neighbourhood.

MS technique is vulnerable to fall into the local maxima, in the case of clutter occlusion [69]. The CAMShift (CMS) is a tracking technique which is an improved form of the MS method. The MS algorithm functions on the Color Probability Distributions (CPDs) & the CAMShift (CMS) is an improved form of the MS to contract with the dynamical changes of the CPDs. To track colour objects in the video frame orders, the color image data has to be signified as a possibility distribution [72,

73]. To complete this, Bradski used the color histograms in his study [72]. The Coupled CMS algorithm as designated in the Bradski's study [72] (& as the reviewed François's study [73]) is established in the real-time head tracking application, which is the part of the Intel OpenCV Computer Vision library [74].

There are numerous papers which address the problem of the moving object tracking from image arrangements captured from the stationary cameras. Based on earlier work on the video segmentation using the joint space-time-range mean shift, authors spread the system to permit the tracking of the moving objects. Large movements of the pdf modes in the consecutive image frames are demoralized for the tracking [7-13].

The optical flow the optic flow is the pattern of ostensible motion of the objects, surfaces, & edges in the visual scene produced by relative motion between an observer (an eye a camera) & the scene. A research [1] suggests a novel technique to precisely estimation the optical-flow for totally automated tracking of the moving-objects within a video streams. This technique contains of an occlusion-killer algorithm & the template matching using a segmented regions, which is achieved by the skip-labelling algorithm. Takaya, K. [2] presents a presentation of the classical technique of active contour (snakes) to real time video environment with the optical flow based technique, where tracking the video object is a primary task. The greedy technique that reduces the energy function to inform the snake points, was collective with optical flow technique to cope with the condition that the object of attention moves also far outside the reach of an edge sensor.

Many author presents researches by combining optical flow with other efficient schemes like gradient snake, state dynamics, kirsch operator, Simultaneous estimation approach as in [3], [4], [5] and [6]. Discrepancy the optical flow approaches are broadly used within the computer vision communal. They are confidential as being whichever local, as in the Lucas-Kanade technique, global, as in the Horn-Schunck method. As the physical subtleties of an object is integrally coupled into behaviour of its image in video stream, in [6], author use such dynamic parameter info in scheming the optical flow when tracking the moving object using the video stream. Certainly, author use an improved error function in minimization that comprises the physical parameter info.

### 3. ACTIVE CONTOUR MODEL, LAYER AND REGION BASED APPROACHES

Another image segmentation approach is active contour models (ACMs), which are in scope of edge-based segmentation and used on object tracking process, as well. A snake is an energy minimizing spline, which is a kind of active contour model. The snake's energy is based on its shape and location within the image. Desired image properties are usually relevant to the local minima of this energy. ACMs, which was suggested by Kass et. al. [75] in 1987 is also known as Snake method.

A snake can be considered as a group of control points (or snaxel) connected to each other and can easily be deformed under applied force. According to the study of Dagher and Tom [76], the situation, in which a snake works the most abundant, is the situation where the points are at the adequate distance and the situation, in which the initial position's coordinates are controlled.

In [14], the new system of semantic video object tracking using the backward region-based organization. It contains of 5 elementary stages: region pre-processing, the region-based motion estimation, region extraction, region classification & the region post-processing. Promising investigational results have exposed the solid presentation of this generic tracking system with the pixel-wise correctness. Another method is to start with the rough user input VO definition. In [15] it then associates each frame's region segmentation & the motion estimation results to build the objects of attention for the temporal tracking this object along the time. An active outline model based algorithm is active to further fine-tune the object's outline so as to extract the correct object boundary.

Automatic detection of the Region of Interests (ROIs) in the complex image video, like an angiogram the endoscopic neurosurgery video, is a serious task in numerous medical image & the video processing applications. An author in [16] finds application of region based approach in such kind of problems. Some of the region based approaches are developed by combining it with techniques like Bayes classifier, wavelets etc. as in [17], [18] and [19]. Many studies are done in enhancing the traditional active contour models in order to enhance results as in [20], [21], [22], [23], & in [24]. As in [20], current a novel moving object detection & the tracking system that strongly fuses infrared & noticeable the video within the level set framework. Author too present the idea of flux tensor as a simplification of 3D structure tensor for fast & dependable motion detection without the Eigen- decomposition. A new association of Active Contour Model (ACM) with the Unscented Kalman Filter (UKF) to track deformable objects in a video sequence present in [23]. A novel model of active contour, VSnares, is presented as a method of segmentation in [24] framework. Associated to the actual snake's algorithm, the semiautomatic video object segmentation with VSnares algorithm occasioned in the enhanced presentation in the terms of video object shape alteration

### 4. STOCHASTIC AND STATICAL APPROACHES

Statistic and stochastic theory is extensively used in motion segmentation field. In fact, the motion segmentation can be seen like a classification problem where every pixel has to be confidential as a background / foreground. The statistical methods can be further separated dependent on the framework used. The common frameworks are the Maximum A posteriori Probability (MAP), the Particle Filter (PF) & the Expectation Maximization (EM). Statistical methods offer the general tool that can be used in very dissimilar way contingent on specific method. The MAP is often used in the mixture with other methods. For e.g., in [77] is mutual with the Probabilistic Data Association Filter. In [78] the MAP is used together with the level sets including motion info. In [79] the MAP frame work is used to association & activity the

interdependence between the motion estimation, segmentation & the super resolution. An author in [25] improve the compressed domain method. At 1st, the motion vectors are collected over an insufficient frames to improve the motion info, which are further spatially interposed to get dense the motion vectors. The last segmentation, using dense motion vectors, is attained by applying the Expectation Maximization (EM) algorithm. A block-based affine clustering technique is proposed for defining the no. of suitable motion models to be used for EM step & the segmented objects are temporally followed to acquire the video objects. There are numerous applications of the EM algorithm & particle filter as in [25, 26, 27, 28, 29, 30 and 31].

The stochastic method, which is based on modeling of an image as an understanding of the random process. Typically, it is expected that the image intensity originates from the Markov Random Field (MRF) &, consequently, contents properties of locality & stationary, i.e. every pixel is only connected to the small set of neighboring pixels & the dissimilar regions of image are professed similar. The Markov random field (MRF) are comeunder this classification. A revolutionary work on retrieval of the plane image geometry is due to [80]. Their algorithm starts with the detection of boundaries of an image objects. The subsequent step is the documentation of the blocked & blocking objects. To this aim, [80] had the glowing idea to imitator the natural capacity of the human vision to whole incompletely occluded objects, supposed a modal conclusion process defined & studied by Gestalt school of psychology & mainly [81]. The theory is applied to a specific model of MRF recognition presented in [82].

Author in [32], propose a new video object tracking method based on the kernel density estimation & the Markov random Field (MRF). The interested video objects are 1st segmented by user, & the nonparametric model based on the kernel density approximation is modified for each video object & the residual background, individually. One popular manner in such type of method at the object detection stage, spatial features of the object are removed by the wavelet transform, according to the frame difference, moving target is determined. In order to efficiently exploit the temporal motion info, the Markov Random Field prior possibility model & the observation field model are recognized, taking benefit of the Bayesian criterion, location of the object in consecutive frame is assessed, & it donates to the correct space object detection [33]. Some author use the Markov random filed in a modified way as done in the literatures [34] [35] and [36].

## 5. DETERMINISTIC AND FEATURE BASED APPROACHES

The deterministic method, whose foremost purpose is to improve the geometry of image, Superiority of skeleton is that it contains useful information of the shape, especially topological and structural information. To have skeleton of a shape, first boundary or edge of the shape is extracted using edge detection algorithms [95] and then its skeleton is generated by skeleton extraction methods [96] Medial axis is a type of skeleton that is defined as the locus of centers of maximal disks that fit within the shape [97]. In [98] present an algorithm for mechanically estimating the subject's skeletal structure from the optical motion capture the data without using any a priori skeletal model. In [99] other researchers have worked on skeleton fitting techniques for use with optical motion capture data. In [100] define partly the automatic method for concluding the skeletons from motion. They resolve for the joint positions by finding the center of revolution in inboard frame for the markers on outboard segment of every joint. The technique of [101], works with the distance restraints although they still trust on the rotation estimates. A rare specific e.g. of approaches for concluding info about a human subject's skeletal structure from motions of bone / skin mounted markers can be establish in [102]. In [103] they have available a survey of the standardization by parameter approximation for robotic devices. Few edge and junction based



researches also done in the field [106, 107]. The approaches that use the edge-based feature kind excerpt the edge map of an image & recognize the topographies of object in terms of the edges. Some e.g. contain [83, 84, 88-101]. Using the edges as features is beneficial over other features due to numerous reasons. As discoursed in [88], they are basically invariant to the illumination conditions & the variations in the objects' colors & the textures. They too signify the object boundaries well & represent the data competently in the large spatial amount of the images.

The other prevalent feature type is the patch based feature type, which uses appearance as cues. This feature has been in use since more than two decades [108], and edge-based features are relatively new in comparison to it. Moravec [108] looked for local maxima of minimum intensity gradients, which he called corners and selected a patch around these corners. His work was improved by Harris [109], which made the new detector less sensitive to noise, edges, and anisotropic nature of the corners proposed in [108]. In this feature type, there are two main variations:

- Patches of rectangular shapes that contain the characteristic boundaries describing the features of the objects [110-115]. Usually, these features are referred to as the local features.
- Irregular patches in which, each patch is homogeneous in terms of intensity or texture and the change in these features are characterized by the boundary of the patches. These features are commonly called the region-based features.

## 6. IMAGE DIFFERENCE AND SUBTRACTION BASED APPROACHES

The object detection can be attained by building the representation of scene known as the background model & then finding the deviations from model for every incoming frame. Any important change in an image region from background model indicates the moving object. Typically, the connected component algorithm is applied to acquire the connected regions corresponding to objects. This process is mentioned to as background subtraction [116]. At each new frame foreground pixels are detected by subtracting intensity values from background and filtering absolute value of differences with dynamic threshold per pixel [117]. The threshold and reference background are updated using foreground pixel information. Eigen background subtraction [118] proposed by Oliver, et al. It presents that an Eigen space model for moving object segmentation. In this method, dimensionality of the space constructed from sample images is reduced by the help of Principal Component Analysis (PCA). It is proposed that the reduced space after PCA should represent only the static parts of the scene, remaining moving objects, if an image is projected on this space [117].

Temporal differencing technique uses the pixel-wise modification between 2-3 consecutive frames in the video imagery to remove the moving regions. It is an extremely adaptive method to dynamic scene changes though, it fails to remove all applicable pixels of the foreground object particularly when the object has unbroken texture moves slowly [119]. A technique for tracking the multiple objects in video arrangements based on the background removal & SIFT feature identical where the camera is stable & the input video orders are real time / the self-captured. Object is distinguished automatically by the background subtraction, then effective tracking is achieved by perceiving the motion & the SIFT feature matching of detected object is described in [37]. Other researches in the including [38, 39, 40, 41, and 42] for applications like Multiple human object tracking, video background estimation, a real-time visual tracking system for dense traffic intersections, real time tracking for surveillance and security system, etcetera.

## 7. SUPERVISED AND SEMI-SUPERVISED LEARNING BASED APPROACHES

Object detection can be performed by learning different object views automatically from a set of examples by means of a supervised learning mechanism. Learning of different object views waives the requirement of storing a complete set of templates. Given a set of learning examples, supervised learning methods generate a function that maps inputs to desired outputs.

A standard formulation of supervised and semi-supervised learning is the classification problem where the learner approximates the behavior of a function by generating an output in the form of either a continuous value, which is called regression, or a class label, which is called classification. In context of object detection, the learning examples are composed of pairs of object features and an associated object class where both of these quantities are manually defined. This classifiers includes Support Vector Machines, Neural Networks, and Adaptive Boosting etcetera.

For detection & tracking of the specific objects in the knowledge- based framework. The arrangement in [43] uses a managed learning technique: the support vector machines. Both problems, detection & tracking, are resolved by a common method: objects are situated in the video sequences by the SVM classifier.

A current controlling trend in the tracking known as the tracking- by-detection uses on-line classifiers in instruction to redetect objects over the succeeding frames. The semi-supervised learning permits for the incorporating priors & is more vigorous in case of obstructions while the multiple-instance learning resolutions the uncertainties where to take the positive informs during tracking. In this work, [44] an algorithm of on-line semi-supervised learning which is capable to association both of these methods into the coherent framework. There are few learning based methods are come in this category including neural networks, support vector machine [45, 46, 47, 48, 49, 50, 51 and 52].

## 8. CONCLUSIONS & DISCUSSIONS

In this editorial, we present a widespread survey of video segmentation and object tracking methods and also give a short- lived review of related matters. We distribute the tracking & the segmentation approaches into 6 groups based on use of the object representations, specifically, approaches founding the point correspondence, approaches using the primitive geometric models, & the approaches using contour evolution. Note that completely these classes need the object detection at some point. For example, the point trackers need detection in each frame, while the geometric region the contours-based trackers need detection only when object 1st appears in the scene. Identifying the significance of the object detection for tracking systems, we comprise a short discussion on the popular object detection approaches. We offer the detailed summaries of object trackers, counting conversation on the object representations, motion models, & parameter estimation schemes active by the tracking algorithms. Furthermore, we define the situation of use, degree of the applicability, evaluation criteria, & the qualitative list of tracking algorithms. We have self-assurance that, this article, the

1st survey on the object tracking with an amusing bibliography Content, can give appreciated insight into this significant research topic & encourage the novel research.

One experiment in tracking is to improve the algorithms for a tracking objects in unconstrained videos, for e.g., the videos attained from the broadcast news networks the home videos. These videos are noisy, unstructured, compressed, & classically contain edited clips developed by moving cameras from the multiple views. Another connected video domain is of official & an informal meetings. These videos typically comprise the multiple people in a small field of view. Thus,

there is simple occlusion, & people are only incompletely visible. One stimulating solution in this context is to employment audio in the addition to video for the object tracking. There are certain approaches being established for estimating point of location of an audio source, for e.g., a mouth of person, based on 4 6 microphones. This audio-based localization of speaker offers the additional info which then can be used in combination with the video-based tracker to resolve the problems such as severe occlusion. Complete, we trust that additional sources of info, in particular prior & the contextual info should be exploited when imaginable to attune the tracker to specific consequence in which it is used. An upright method to integrate these disparate sources of info will result in an overall tracker that can be working with success in a variability of applications.

## 9. REFERENCES

1. Hiraiwa, A., Accurate estimation of optical flow for fully automated tracking of moving-objects within video streams, Circuits and Systems, 1999. ISCAS '99. Proceedings of the 1999 IEEE International Symposium
2. Takaya, K. Tracking a video object with the active contour (snake) predicted by the optical flow, Electrical and Computer Engineering, 2008. CCECE 2008. Canadian Conference
3. Jayabalan, E., Non Rigid Object Tracking in Aerial Videos by Combined Snake and Optical flow Technique, Computer Graphics, Imaging and Visualisation, 2007. CGIV'07
4. Bauer, N.J., Object focused simultaneous estimation of optical flow and state dynamics, Intelligent Sensors, Sensor Networks and Information Processing, 2008. ISSNIP 2008. International Conference
5. Ping Gao, Moving object detection based on kirsch operator combined with Optical Flow, Image Analysis and Signal Processing (IASP), 2010 International Conference
6. Pathirana, P.N., Simultaneous estimation of optical flow and object state: A modified approach to optical flow calculation, Networking, Sensing and Control, 2007 IEEE International Conference
7. Tiesheng Wang, Moving Object Tracking from Videos Based on Enhanced Space-Time-Range Mean Shift and Motion Consistency, Multimedia and Expo, 2007 IEEE International Conference
8. Khan, Z.H., Joint particle filters and multi-mode anisotropic mean shift for robust tracking of video objects with partitioned areas, Image Processing (ICIP), 2009 16th IEEE International Conference
9. Yuanyuan Jia, Real-time integrated multi-object detection and tracking in video sequences using detection and mean shift based particle filters, Web Society (SWS), 2010 IEEE 2nd Symposium
10. Sung-Mo Park, Object tracking in MPEG compressed video using mean-shift algorithm, Information, Communications and Signal Processing, 2003 and Fourth Pacific Rim Conference on Multimedia. Proceedings of the 2003 Joint Conference of the Fourth International Conference
11. Khan, Z.H., Joint anisotropic mean shift and consensus point feature correspondences for object tracking in video, Multimedia and Expo, 2009. ICME 2009. IEEE International Conference
12. Yilmaz A., Object Tracking by Asymmetric Kernel Mean Shift with Automatic Scale and Orientation Selection, Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference

13. Da Tang, Combining Mean-Shift and Particle Filter for Object Tracking, Image and Graphics (ICIG), 2011 Sixth International Conference
14. Chuang Gu, Semantic video object tracking using region-based classification, Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference
15. Dongxiang Xu, An accurate region based object tracking for video sequences, Multimedia Signal Processing, 1999 IEEE 3rd Workshop
16. Bing Liu, Automatic Detection of Region of Interest Based on Object Tracking in Neurosurgical Video, Engineering in Medicine and Biology Society, 2005. IEEE- EMBS 2005. 27th Annual International Conference
17. Mezaris, V., Video object segmentation using Bayes- based temporal tracking and trajectory-based region merging, Circuits and Systems for Video Technology, IEEE Transactions
18. Khraief, C, Unsupervised video objects detection and tracking using region based level-set, Multimedia Computing and Systems (ICMCS), 2012 International Conference
19. Kai-Tat Fung, Region-based object tracking for multipoint video conferencing using wavelet transform, Consumer Electronics, 2001. ICCE. International Conference
20. Bunyak, F., Geodesic Active Contour Based Fusion of Visible and Infrared Video for Persistent Object Tracking, Applications of Computer Vision, 2007. WACV '07. IEEE Workshop
21. Allili, M.S., A Robust Video Object Tracking by Using Active Contours, Computer Vision and Pattern Recognition Workshop, 2006. CVPRW '06. Conference
22. Castaud, M., Tracking video objects using active contours, Motion and Video Computing, 2002. Proceedings Workshop
23. Messaoudi, Z., Tracking objects in video sequence using active contour models and unscented Kalman filter, Systems, Signal Processing and their Applications (WOSSPA), 2011 7th International Workshop
24. Shijun Sun, Semiautomatic video object segmentation using VSnares, Circuits and Systems for Video Technology, IEEE Transactions
25. Babu, R.V., Video object segmentation: a compressed domain approach, Circuits and Systems for Video Technology, IEEE Transactions
26. Hai Tao, Object tracking with Bayesian estimation of dynamic layer representations, Pattern Analysis and Machine Intelligence, IEEE Transactions
27. Qi Hao, Multiple Human Tracking and Identification with Wireless Distributed Pyroelectric Sensor Systems, Systems Journal, IEEE
28. Marlow, S., Supervised object segmentation and tracking for MPEG-4 VOP generations, Pattern Recognition, 2000. Proceedings. 15th International Conference
29. Beiji Zou ; Peng, Xiaoning ; Han, Liqin, Particle Filter with Multiple Motion Models for Object Tracking in Diving Video Sequences, Image and Signal Processing, 2008. CISP'08. Congress

30. Tao Zhang, Modified particle filter for object tracking in low frame rate video, Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference
31. Kumar, T.Senthil, Object detection and tracking in video using particle filter, Computing Communication & Networking Technologies (ICCCNT), 2012 Third International Conference
32. Zhi Liu, A Novel Video Object Tracking Approach Based on Kernel Density Estimation and Markov Random Field, Image Processing, 2007. ICIP 2007. IEEE International Conference
33. Xuchao Li, Object Detection and Tracking Using Spatiotemporal Wavelet Domain Markov Random Field in Video, Computational Intelligence and Security, 2009. CIS'09. International Conference
34. Khatoonabadi, S. H., Video Object Tracking in the Compressed Domain Using Spatio-Temporal Markov Random Fields, Image Processing, IEEE Transactions
35. Lefevre, S., A new way to use hidden Markov models for object tracking in video sequences, Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference
36. Guler, S., A Video Event Detection and Mining Framework, Computer Vision and Pattern Recognition Workshop, 2003. CVPRW '03. Conference
37. Rahman, M.S., Multi-object Tracking in Video Sequences Based on Background Subtraction and SIFT Feature Matching, Computer Sciences and Convergence Information Technology, 2009. ICCIT '09. Fourth International Conference
38. Dharamadhat, T., Tracking object in video pictures based on background subtraction and image matching, Robotics and Biomimetics, 2008. ROBIO 2008. IEEE International Conference
39. Susheel Kumar, K., Multiple Cameras Using Real Time Object Tracking for Surveillance and Security System, Emerging Trends in Engineering and Technology (ICETET), 2010 3rd International Conference
40. Saravanakumar, S., Multiple human object tracking using background subtraction and shadow removal techniques, Signal and Image Processing (ICSIP), 2010 International Conference
41. Scott, J., Kalman filter based video background estimation, Applied Imagery Pattern Recognition Workshop (AIPRW), 2009 IEEE
42. Rostamianfar, O., Visual Tracking System for Dense Traffic Intersections, Electrical and Computer Engineering, 2006. CCECE '06. Canadian Conference
43. Carminati, L., Knowledge-Based Supervised Learning Methods in a Classical Problem of Video Object Tracking, Image Processing, 2006 IEEE International Conference
44. Zeisl, B., On-line semi-supervised multiple-instance boosting, Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference
45. Nan Jiang, Learning Adaptive Metric for Robust Visual Tracking, Image Processing, IEEE Transactions
46. Han, T.X., A Drifting-proof Framework for Tracking and Online Appearance Learning, Applications of Computer

- Vision, 2007. WACV '07. IEEE Workshop
47. Cerman, L., Tracking with context as a semi-supervised learning and labeling problem, Pattern Recognition (ICPR), 2012 21st International Conference
  48. Babenko, B., Robust Object Tracking with Online Multiple Instance Learning, Pattern Analysis and Machine Intelligence, IEEE Transactions
  49. Guorong Li, Treat samples differently: Object tracking with semi-supervised online CovBoost, Computer Vision (ICCV), 2011 IEEE International Conference
  50. Fang Tan, A method for robust recognition and tracking of multiple objects, Communications, Circuits and Systems, 2009. ICCAS 2009. International Conference
  51. Takaya, K., Tracking moving objects in a video sequence by the neural network trained for motion vectors, Communications, Computers and signal Processing, 2005. PACRIM. 2005 IEEE Pacific Rim Conference
  52. Doulamis, N., Video object segmentation and tracking in stereo sequences using adaptable neural networks, Image Processing, 2003. ICIP 2003. Proceedings. 2003 International Conference
  53. M. Teklap, Digital Video Processing. Prentice Hall, NJ, 1995.
  54. P. Salembier and F. Marques, "Region based representation of image and video segmentation tools for multimedia services," IEEE Trans. Circuit systems and video Technology, vol. 9, No. 8, pp. 1147-1169, Dec. 1999.
  55. E. Y. Kim, S. H. Park and H. J. Kim, "A Genetic Algorithm-based segmentation of Random Field Modeled images," IEEE Signal processing letters, vol. 7, No. 11, pp. 301-303, Nov. 2000.
  56. S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 6, No. 6, pp. 721-741, Nov. 1984.
  57. R. O. Hinds and T. N. Pappas, "An Adaptive Clustering algorithm for Segmentation of Video Sequences," Proc. of International Conference on Acoustics, speech and signal Processing, ICASSP, vol. 4, pp. 2427-2430, May. 1995.
  58. E. Y. Kim and K. Jung, "Genetic Algorithms for video Segmentation," Pattern Recognition, vol. 38, No. 1, pp. 59-73, 2005.
  59. E. Y. Kim and S. H. Park, "Automatic video Segmentation using genetic algorithms," Pattern Recognition Letters, vol. 27, No. 11, pp. 1252-1265, Aug. 2006.
  60. Stan Z. Li, Markov field modeling in image analysis. Springer: Japan, 2001.
  61. J. Besag, "on the statistical analysis of dirty pictures," Journal of Royal Statistical Society Series B (Methodological), vol. 48, No. 3, pp. 259-302, 1986.
  62. A.L. Bovik, Image and Video Processing. Academic Press, New York, 2000.
  63. G. K. Wu and T. R. Reed, "Image sequence processing using spatiotemporal segmentation," IEEE Trans. on circuits and systems for video Technology, vol. 9, No. 5, pp. 798-807, Aug. 1999.

64. S. W. Hwang, E. Y. Kim, S. H. Park and H. J. Kim, "Object Extraction and Tracking using Genetic Algorithms," Proc. of International Conference on Image Processing, Thessaloniki, Greece, vol.2, pp. 383-386, Oct. 2001.
65. S. D. Babacan and T. N. Pappas, "Spatiotemporal algorithm for joint video segmentation and foreground detection," Proc. EUSIPCO, Florence, Italy, Sep. 2006.
66. S. D. Babacan and T. N. Pappas, "Spatiotemporal Algorithm for Background Subtraction," Proc. of IEEE International Conf. on Acoustics, Speech, and Signal Processing, ICASSP 07, Hawaii, USA, pp. 1065-1068, AprilX2007.
67. [67] S. S. Huang and L. Fu," Region-level motion-based background modeling and subtraction using MRFs," IEEE Transactions on image Processing, vol.16, No. 5, pp.1446-1456, May. 2007.
68. Comaniciu, D., Meer, P., "Mean shift: A robust approach toward feature space analysis", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 5, pp. 603–619, 2002.
69. Shan, C., Tan, T., Wei, Y., "Real-time hand tracking using a mean shift embedded particle filter", Pattern Recognition, Vol. 40, No. 7, pp. 1958-1970, 2007.
70. Yilmaz, A., Javed, O., Shah, M., "Object tracking: A survey", ACM Computing Surveys, Vol. 38, No. 4, Article No. 13 (Dec. 2006), 45 pages, doi:10.1145/1177352.1177355, 2006.
71. Comaniciu, D., Meer, P., "Mean Shift Analysis and Applications", IEEE International Conference Computer Vision (ICCV'99), Kerkyra, Greece, pp. 1197-1203, 1999.
72. Bradski, G. "Computer Vision Face Tracking for Use in a Perceptual User Interface", In Intel Technology Journal,([http://developer.intel.com/technology/itj/q21998/articles/art\\_2.htm](http://developer.intel.com/technology/itj/q21998/articles/art_2.htm), (Q2 1998).
73. François, A. R. J., "CAMSHIFT Tracker Design Experiments with Intel OpenCV and SAI", IRIS Technical Report IRIS-04-423, University of Southern California, Los Angeles, 2004.
74. Bradski, G., Kaehler, A., "Learning OpenCV: Computer Vision with the OpenCV Library", O'Reilly Media, Inc. Publication, 1005 Gravenstein Highway North, Sebastopol, CA 95472, 2008.
75. Kass, M., Witkin, A., and Terzopoulos, D., "Snakes: active contour models", International Journal of Computer Vision, Vol. 1, No. 4, pp. 321–331, 1988.
76. Dagher, I., Tom, K. E., "WaterBalloons: A hybrid watershed Balloon Snake segmentation", Image and Vision Computing, Vol. 26, pp. 905–912, doi:10.1016/j.imavis.2007.10.010, 2008.
77. C. Rasmussen and G. D. Hager, "Probabilistic data association methods for tracking complex visual objects," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 6, pp. 560–576, 2001.
78. D. Cremers and S. Soatto, "Motion competition: A variational approach to piecewise parametric motion segmentation," International Journal of Computer Vision, vol. 62, no. 3, pp. 249–265, May 2005.
79. H. Shen, L. Zhang, B. Huang, and P. Li, "A map approach for joint motion estimation, segmentation, and super resolution," IEEE Transactions on Image Processing, vol. 16, no. 2, pp. 479–490, 2007.
80. Nitzberg, M., Mumford, D. and Shiota, T. (1993), Filtering, Segmentation and Depth, Lecture Notes in Computer

- Science, Vol. 662, Springer-Verlag, Berlin.
81. Kanizsa, G. (1979), Organisation in Vision, New-York: Praeger.
  82. Li, S.Z. (1994). A Markov random field model for object matching under contextual constraints, In proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Seattle, Washington
  83. "John D. Trimmer, 1950, Response of Physical Systems (New York: Wiley), p. 13. 93 IJCSI International Journal of Computer Science Issues, Vol. 7, Issue 6, November 2010 ISSN
  84. Zhiqiang Zheng, Han Wang \*, Analysis of gray level corner detection Eam Khwang Teoh, School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798, Singapore Received 29 May 1998; received in revised form 9 October 1998.
  85. Grimson, W.E.L. (1990). Object Recognition by Computer-The Role of Geometric Constraints. MIT Press, Cambridge, MA.
  86. STANZ, LI, Parameter Estimation for Optimal Object Recognition: Theory and Application, School of Electrical and Electronic Engineering Nanyang Technological University, Singapore 63978
  87. Arthur R., Pope David G. Lowe, David Sarnoff Learning Appearance Models for Object Recognition, , Research Center, CN530, Princeton, J085435300. Dept of Computer Science, University of British Columbia Vancouver, B.C, Canada V6T1Z4.
  88. P Gros. Matching and clustering: Two steps towards automatic object model generation in computer vision. In Proc. AAAI Fall Symp.: Machine Learning in Computer Vision, AAAI Press, 1993.
  89. M. Seibert, A.M Waxman. Adaptive 3-D object recognition from multiple views. IEEE Trans. Patt. Anal. Mach. Intell. 14:107-124, 1992.
  90. Andrea Selinger and Randal C. Nelson, "A Perceptual Grouping Hierarchy for Appearance-Based 3D Object Recognition", Computer Vision and Image Understanding, vol. 76, no. 1, October 1999, pp.83-92. Abstract, gzipped postscript (preprint)
  91. Randal C. Nelson and Andrea Selinger "A Cubist Approach to Object Recognition", International Conference on Computer Vision (ICCV98), Bombay, India, January, 614-621. Abstract, gzipped postscript, also in an extended version with more complete description of the algorithms, and additional experiments.
  92. Randal C. Nelson, "From Visual Homing to Object Recognition", in Visual Navigation, Yiannis Aloimonos, Editor, Lawrence Erlbaum Inc, 1996, 218-250. Abstract,
  93. Randal C. Nelson, "Memory-Based Recognition for 3-D Objects", Proc. ARPA Image Understanding Workshop, Palm Springs CA, February 1996, 1305-1310. Abstract, gzipped postscript
  94. Ashikhmin, M. (2001), Synthesizing Natural Textures, Proc. ACM Symp. Interactive 3D Graphics, Research Triangle Park, USA, March 2001, pp 217-226.
  95. T. Bernier, J.-A. Landry, "A New Method for Representing and Matching Shapes of Natural Objects", Pattern Recognition 36, 2003, pp. 1711-1723.



96. H. Blum, "A Transformation for extracting new descriptors of Shape", in: W. Whaten-Dunn(Ed.), MIT Press, Cambridge, 1967. pp. 362-380.
97. Serge Belongie, Jitendra Malik, Jan Puzicha, "Shape Matching and Object Recognition using Shape Contexts", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.24 , No. 24, pp. 509-522, April 2002.
98. Kirk James F. O'Brien, ,David A. Forsyth, Skeletal Parameter Estimation from Optical Motion Capture Data, Adam University of California, Berkeley
99. J. Costeira and T. Kanade. A multibody factorization method for independently moving-objects. Int. J. Computer Vision, 29(3):159–179, 1998.
100. M.-C. Silaghi, R. Plankers, R. Boulic, P. Fua, and D. Thalmann. Local and global skeleton fitting techniques for optical motion capture. In N. M. Thalmann and D. Thalmann, editors, Modelling and Motion Capture Techniques for Virtual Environments, volume 1537 of Lecture Notes in Artificial Intelligence, pages 26–40, 1998.
- 101.[101] M. Ringer and J. Lasenby. A procedure for automatically estimating model parameters in optical motion capture. In BMVC02, page Motion and Tracking, 2002.
102. J. H. Challis. A procedure for determining rigid body transformation parameters. Journal of Biomechanics, 28(6):733–737, 1995.
103. B. Karan and M. Vukobratović. Calibration and accuracy of manipulation robot models – an overview. Mechanism and Machine Theory, 29(3):479–500, 1992.
104. Hamidreza Zaboli, Shape Recognition by Clustering and Matching of Skeletons, Amirkabir University of Technology, Tehran, Iran Dengsheng Zhang, Guojun Lu , "Review of shape representation and description techniques", Pattern Recognition 37 , 2004, pp. 1-19.
105. Tapas Kanungo, A methodology for quantitative performance evaluation of detection algorithms, Student member, IEEE , M.Y Jaisimha, Student member, IEEE, John Palmer and Robert M. Haralick, Fellow , IEEE
106. E.S Deutschand J.R Fram, A quantitative study of the orientation bias of some edge detector schemes", IEEE Transactions on, Computers, vol C-27, no.3 pp 205-213, March 1978.
107. H. P. Moravec, "Rover visual obstacle avoidance," in Proceedings of the International Joint Conference on Artificial Intelligence, Vancouver, CANADA, 1981, pp. 785-790.
108. C. Harris and M. Stephens, "A combined corner and edge detector," presented at the Alvey Vision Conference, 1988.
109. J. Li and N. M. Allinson, "A comprehensive review of current local features for computer vision," Neurocomputing, vol. 71, pp. 1771-1787, 2008.
110. K. Mikolajczyk and H. Uemura, "Action recognition with motion-appearance vocabulary forest," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2008, pp. 1-8.
111. K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," in Proceedings of the IEEE

- Conference on Computer Vision and Pattern Recognition, 2003, pp. 1615-1630.
112. D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91-110, 2004.
  113. B. Ommer and J. Buhmann, "Learning the Compositional Nature of Visual Object Categories for Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010.
  114. B. Leibe, A. Leonardis, and B. Schiele, "Robust object detection with interleaved categorization and segmentation," *International Journal of Computer Vision*, vol. 77, pp. 259-289, 2008.
  115. Elgammal, A. Duraiswami, R., Hairwood, D., And Davis, L. 2002. Background and foreground modeling using nonparametric kernel density estimation for visual surveillance. *Proceedings of IEEE* 90, 7, 1151–1163.
  116. S. Y. Elhabian, K. M. El-Sayed, "Moving object detection in spatial domain using background removal techniques- state of the art", *Recent patents on computer science*, Vol 1, pp 32-54, Apr, 2008.
  117. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. In *IEEE International Conference on Computer Vision*. 694–699, 1995.
  118. N. Paragios, and R. Deriche. Geodesic active contours and level sets for the detection and tracking of moving objects. *IEEE Trans. Patt. Anal. Mach. Intell.* 22, 3, 266–280, 2000.